MULTIMEDIA UNIVERSITY*

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2022/2023

DSP5614 – PROBABILITY AND STATISTICS

(for diploma students only)

13 JULY 2023
9.00 a.m – 11.00 a.m
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of 5 pages including appendix.
2. Attempt **ALL** questions. The distribution of the marks for each question is given.
3. Please write all your answer in the Answer Booklet provided.
4. Key formulae are given in the Appendix.

QUESTION 1

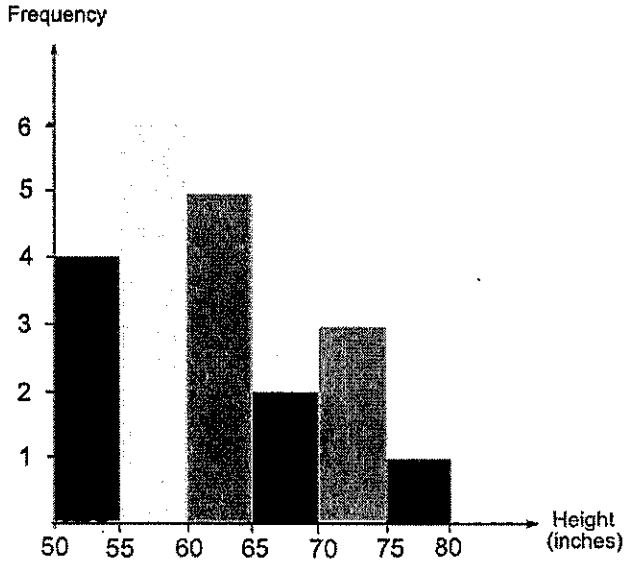
a) Identify the data type for the following. (5 marks)

i. Sarah and her sister go for swimming lessons four times per week.

ii. Type of hair preferences by teenagers.

iii. Temperature of coffee served for each customer at the cafeteria.

b) The following histogram shows the height of students in Art Class.



Based on the diagram above, answer the following:

i. What is the percentage of students that have height more than 70 inches? (3 marks)

ii. Construct a grouped data table. (4 marks)

iii. From the table in (ii), calculate the value of mean. (3 marks)

c) The following is a stem and leaf diagram of the Maths grade of Form 4 students.

Stem	Leaf
5	5
6	0 2 7
7	0 2 6 7 9
8	0 5 5 5 8
9	0 2 5 7 8 9
10	0 0

Key: 5 | 5 = 55

Continue...

i. Calculate the measure of variations. (7 marks)
ii. What is the shape of the distribution? (1 mark)
iii. Find the value of median. (2 marks)

[TOTAL: 25 MARKS]

QUESTION 2

a) Identify whether the following statement is **TRUE / FALSE.** (3 marks)

i. Event A and B are mutually exclusive if $P(A \cap B) = 0$.
ii. Complementary Rule can be used when you want to find the probability of event A or B.
iii. General Multiplication Rule can be used to find the probability if you pick a card from a drawer without replacement.

b) Johan has entered a tennis tournament and wants to practice as often as possible to prepare for it. However, it depends on the weather. Based on past data, the probability that it will rain on any given day is 30%. If it rains, the probability that Johan can practise is 8%. If it does not rain, however, the probability that Johan practise is 90%.

i. Construct a tree diagram for this situation. (6 marks)
ii. What is the probability that Johan will be able to practise for his tournament? (4 marks)

c) The committee of a high school needs to plan for a seating at their annual dinner. The following table displays data for the students involved in the annual dinner this year.

Level	Performing	Not Performing
Junior	35	42
Senior	21	35
Staff	4	13

Find the probability that a random person chosen is
i. the school's staff. (2 marks)
ii. junior who will not be performing. (2 marks)
iii. senior or will be performing. (4 marks)
iv. not performing given that the person is a senior. (4 marks)

[TOTAL: 25 MARKS]

Continued ...

QUESTION 3

a) It is reported that teenagers spend 9.6 hours per week on average with their phone. With the current trend, it is expected to be higher. A random survey was conducted with 150 teenagers on their time spent with phone. Results show sample mean of 9.75 hours and population standard deviation of 5. Conduct a hypothesis test at 1% significance level. (12 marks)

b) There is a claim that about 10% of men and 30% of women traveling on business bring a friend or a spouse. According to a survey by Eastern Hotel, 14% of the 50 men and 20% of the 30 women who are traveling for business purposes brought a friend or spouse. Can Eastern Hotel dispute this claim at 5% significance level? (13 marks)

[TOTAL: 25 MARKS]

QUESTION 4

You want to examine how much money should be allocated for petrol when travelling. The relationship between the total fuel and total distances are shown as following.

Total Fuel (RM)	42	37	35	33	34	35	37	39	38	40
Total Distances (KM)	189	166	157	149	153	158	167	176	171	180

a) Calculate SS_{xx} , SS_{yy} and SS_{xy} based on the information given. (9 marks)

b) Calculate the correlation coefficient, r and interpret the answer. (4 marks)

c) Find the linear regression equation $\hat{y} = \beta_0 + \beta_1 x$. (5 marks)

d) Predict the total distance if the total fuel is RM85. (2 marks)

e) Compute the value of *SST*, *SSR* and *SSE*. (3 marks)

f) Calculate the standard error of the estimate, s_e. (2 marks)

[TOTAL: 25 MARKS]

End of Paper

APPENDIX – KEY FORMULA

	Ungrouped Data	Grouped Data
Mean	$\bar{x} = \frac{\sum x_i}{n}$ where n : sample size	$\bar{x} = \frac{\sum m_i f_i}{\sum f_i}$ where m : class midpoint f : class frequency
Variance	$s^2 = \frac{1}{n-1}\left[\sum {x_i}^2 - \frac{(\sum x_i)^2}{n}\right]$ where n : sample size	$s^2 = \frac{1}{(\sum f_i)-1}\left[\sum {m_i}^2 f_i - \frac{(\sum m_i f_i)^2}{\sum f_i}\right]$ where m : class midpoint f : class frequency

NORMAL AND STANDARD NORMAL PROBABILITY DISTRIBUTION

- z-value (observed value) for an x value : $Z = \frac{x-\mu}{\sigma}$

HYPOTHESIS TESTING

Population	Mean	Proportion
1	$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	$Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ where $\hat{p} = \frac{x}{n}$
2	$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}}$	$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}_p(1-\hat{p}_p)}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $where \hat{p}_p = \frac{x_1 + x_2}{n_1 + n_2}$

LINEAR REGRESSION ANALYSIS

- **Least square regression equation,** $\hat{y} = \beta_0 + \beta_1 x$ where $\boldsymbol{\beta_1} = \frac{S_{xy}}{S_{xx}}$ and $\boldsymbol{\beta_0} = \bar{\boldsymbol{y}} - \boldsymbol{\beta_1}\bar{\boldsymbol{x}}$
- **Correlation coefficient,** $\boldsymbol{r} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$ where

$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$	$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$	$S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$

- **Sum of Square**

$SST = SSR + SSE$	$SST = S_{yy}$	$SSR = \frac{(S_{xy})^2}{S_{xx}}$

- **Test of Significance for Regression Slope**

Standard Error, $s_e = \sqrt{\frac{SSE}{n-2}}$	Regression t-Test $t = \frac{\beta_1}{s_e/\sqrt{Sxx}}$	Correlation t-Test $t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$